

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/525,365  
Source: P4  
Date Processed by STIC: 3/3/06

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 03/03/2006

PATENT APPLICATION: US/10/525,365

TIME: 13:01:12

Input Set : A:\P26794.ST25.txt

Output Set: N:\CRF4\03032006\J525365.raw

3 <110> APPLICANT: MIYAWAKI, Atsushi  
 4 ANDO, Ryoko  
 5 KARASAWA, Satoshi  
 6 MIZUNO, Hideaki  
 8 <120> TITLE OF INVENTION: FLUORESCENT PROTEIN AND CHROMOPROTEIN  
 10 <130> FILE REFERENCE: P26794  
 12 <140> CURRENT APPLICATION NUMBER: US 10/525,365  
 13 <141> CURRENT FILING DATE: 2005-02-23  
 15 <150> PRIOR APPLICATION NUMBER: JP2002/243337  
 16 <151> PRIOR FILING DATE: 2002-08-23  
 18 <150> PRIOR APPLICATION NUMBER: JP2002/243338  
 19 <151> PRIOR FILING DATE: 2002-08-23  
 21 <150> PRIOR APPLICATION NUMBER: JP2002/274266  
 22 <151> PRIOR FILING DATE: 2002-09-20  
 24 <150> PRIOR APPLICATION NUMBER: JP2002/280118  
 25 <151> PRIOR FILING DATE: 2002-09-26  
 27 <150> PRIOR APPLICATION NUMBER: National Stage of PCT/JP2003/010628  
 28 <151> PRIOR FILING DATE: 2003-08-22  
 30 <160> NUMBER OF SEQ ID NOS: 34  
 32 <170> SOFTWARE: PatentIn version 3.3  
 34 <210> SEQ ID NO: 1  
 35 <211> LENGTH: 229  
 36 <212> TYPE: PRT  
 37 <213> ORGANISM: Anthopleura inornata  
 39 <400> SEQUENCE: 1  
 41 Met Ala Thr Leu Val Lys Glu Thr Met Arg Ile Lys Met Ser Met Glu  
 42 1 5 10 15  
 45 Gly Thr Val Asn Gly His His Phe Lys Cys Glu Gly Gln Gly Glu Gly  
 46 20 25 30  
 49 Lys Pro Phe Glu Gly Tyr Gln Val Glu Lys Ile Arg Val Thr Glu Gly  
 50 35 40 45  
 53 Gly Pro Leu Pro Phe Ala Tyr Asp Thr Leu Thr Pro Cys Trp Met Tyr  
 54 50 55 60  
 57 Gly Ser Lys Thr Phe Ile Lys His Thr Ser Gly Ile Pro Asp Tyr Phe  
 58 65 70 75 80  
 61 Lys Glu Ser Leu Pro Glu Gly Phe Thr Trp Glu Arg Thr Gln Ile Tyr  
 62 85 90 95  
 65 Glu Asp Gly Gly Cys Leu Thr Ile His Gln Asp Thr Ser Met Gln Gly  
 66 100 105 110  
 69 Asp Cys Phe Ile Phe Lys Ile Lys Val Ile Gly Thr Asn Phe Pro Ala  
 70 115 120 125  
 73 Asn Gly Pro Val Met Gln Lys Lys Thr Ala Gly Trp Glu Pro Cys Val  
 74 130 135 140

P. 6

## RAW SEQUENCE LISTING

DATE: 03/03/2006

PATENT APPLICATION: US/10/525,365

TIME: 13:01:12

Input Set : A:\P26794.ST25.txt

Output Set: N:\CRF4\03032006\J525365.raw

```

77 Glu Met Leu Tyr Pro Arg Ala Gly Val Leu Cys Gly Gln Ser Leu Met
78 145                      150                      155                      160
81 Ala Leu Lys Cys Lys Asp Gly Asn His Leu Thr Cys His Leu Arg Thr
82                      165                      170                      175
85 Thr Tyr Arg Ser Arg Lys Ala Gly Gln Lys Met Pro Glu Phe His Phe
86                      180                      185                      190
89 Gly Asp His Arg Ile Glu Ile Leu Lys Glu Glu Glu Gln Gly Met Arg
90                      195                      200                      205
93 Ile Glu Gln Tyr Glu Ala Ala Val Ala Arg Tyr Cys Glu Ala Pro Ser
94                      210                      215                      220
97 Arg Leu Gly His His
98 225

```

101 &lt;210&gt; SEQ ID NO: 2

102 &lt;211&gt; LENGTH: 690

103 &lt;212&gt; TYPE: DNA

104 &lt;213&gt; ORGANISM: Anthopleura inornata

106 &lt;400&gt; SEQUENCE: 2

```

107 atggctacct tggtaaaga aactatgcgc atcaagatga gtatggaagg gacggttaat      60
109 ggacaccacg tcaagtgtga aggacaagga gagggcaagc cttttgaagg ttaccaggctc      120
111 gaaaagatta gagtactga aggaggtccg ctaccctttg cgtacgatac tttagacacct      180
113 tgctggatgt atggaagtaa aaccttcacg aagcatacat caggaattcc cgattacttc      240
115 aaggagtctc ttcctgaagg ctttacttgg gaaagaacgc aaatctacga ggatggaggc      300
117 tgtcttacta ttcaccagga cacaagcatg caggagatt gttttatatt caagataaaa      360
119 gtcattggta ccaactttcc tgccaatggt cccgtgatgc agaagaaaac agcaggatgg      420
121 gagccatgcg ttgagatgct ttatcctcgt gccggtgtct tgtgtggaca gtcgttgatg      480
123 gccctgaaat gcaaggatgg caaccacctg acgtgccatc tgcgaactac ctacagggtc      540
125 agaaaggcag gacaaaaaat gccagagttc catttcgggg atcatcgat tgagatcctg      600
127 aaggaagaag aacaaggcat gcgtattgaa caatacgagg cagcggtggc gaggtactgc      660
129 gaagctccat ccaggcttgg acatcactaa
130                                     690

```

132 &lt;210&gt; SEQ ID NO: 3

133 &lt;211&gt; LENGTH: 229

134 &lt;212&gt; TYPE: PRT

135 &lt;213&gt; ORGANISM: Anthopleura inornata

137 &lt;400&gt; SEQUENCE: 3

```

139 Met Ala Thr Leu Val Lys Glu Thr Met Arg Ile Lys Met Ser Met Glu
140 1                      5                      10                      15
143 Gly Thr Val Asn Gly His His Phe Lys Cys Glu Gly Gln Gly Glu Gly
144                      20                      25                      30
147 Lys Pro Phe Glu Gly Tyr Gln Val Glu Lys Ile Arg Val Thr Glu Gly
148                      35                      40                      45
151 Gly Pro Leu Pro Phe Ala Tyr Asp Ile Leu Ala Pro Cys Cys Ser Tyr
152                      50                      55                      60
155 Gly Ser Lys Thr Phe Ile Lys His Val Ser Gly Ile Pro Asp Tyr Phe
156 65                      70                      75                      80
159 Lys Glu Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Gln Ile Tyr
160                      85                      90                      95
163 Glu Asp Gly Gly Ser Leu Ser Ile His Gln Asp Thr Ser Leu Gln Gly
164                      100                     105                     110
167 Asp Cys Phe Ile Tyr Lys Ile Lys Val Ile Gly Thr Asn Phe Pro Ala

```

## RAW SEQUENCE LISTING

DATE: 03/03/2006

PATENT APPLICATION: US/10/525,365

TIME: 13:01:12

Input Set : A:\P26794.ST25.txt

Output Set: N:\CRF4\03032006\J525365.raw

```

168          115          120          125
171 Asn Gly Pro Val Met Gln Lys Lys Thr Ala Gly Trp Glu Pro Cys Val
172          130          135          140
175 Glu Met Leu Tyr Pro Arg Ala Gly Val Leu Cys Gly Gln Ser Leu Met
176 145          150          155          160
179 Ala Leu Lys Cys Lys Asp Gly Asn His Leu Thr Cys His Leu Arg Thr
180          165          170          175
183 Thr Tyr Arg Ser Arg Lys Ala Gly Gln Lys Met Pro Glu Phe His Phe
184          180          185          190
187 Gly Asp His Arg Ile Glu Ile Leu Lys Glu Glu Glu Gln Gly Met Arg
188          195          200          205
191 Ile Glu Gln Tyr Glu Ala Ala Val Ala Arg Tyr Cys Glu Ala Pro Ser
192          210          215          220
195 Arg Leu Gly His His
196 225
199 <210> SEQ ID NO: 4
200 <211> LENGTH: 690
201 <212> TYPE: DNA
202 <213> ORGANISM: Anthopleura inornata
204 <400> SEQUENCE: 4
205 atggctacct tggttaaaga aactatgcgc atcaagatga gtatggaagg gacggttaat 60
207 ggacaccact tcaagtgtga aggacaagga gagggcaagc cttttgaagg ttaccaggtc 120
209 gaaaagatta gagttactga aggaggtccg ctaccctttg cgtacgatat tttggcacct 180
211 tgctgctcgt atggaagtaa aaccttcac cagcatgtct cgggaatccc cgattacttc 240
213 aaggagtcct tccctgaagg ctttacttgg gaaagaacgc aaatctacga ggatggaggc 300
215 tctctttcta ttcaccagga cacaagcctg caggagatt gttttattta caagatcaaa 360
217 gtcattggca ccaactttcc tgccaatggt cccgtgatgc agaagaaaac agcaggatgg 420
219 gagccatgcg ttgagatgct ttatcctcgt gccggtgtct tgtgtggaca gtcgttgatg 480
221 gccctgaaat gcaaggatgg caaccacctg acgtgccatc tgcgaaactac ctacaggtcc 540
223 agaaaggcag gacaaaaaat gccagagttc catttcgggg atcatcgat tgagatcctg 600
225 aaggaagaag aacaaggcat gcgtattgaa caatacgagg cagcgggtggc gaggtactgc 660
227 gaagctccat ccaggcttgg acatcactaa 690
230 <210> SEQ ID NO: 5
231 <211> LENGTH: 225
232 <212> TYPE: PRT
233 <213> ORGANISM: Trachyphyllia geoffroyi
235 <400> SEQUENCE: 5
237 Met Ser Leu Ile Lys Pro Glu Met Lys Ile Lys Leu Leu Met Glu Gly
238 1          5          10          15
241 Asn Val Asn Gly His Gln Phe Val Ile Glu Gly Asp Gly Lys Gly His
242          20          25          30
245 Pro Phe Glu Gly Lys Gln Ser Met Asp Leu Val Val Lys Glu Gly Ala
246          35          40          45
249 Pro Leu Pro Phe Ala Tyr Asp Ile Leu Thr Thr Ala Phe His Tyr Gly
250          50          55          60
253 Asn Arg Val Phe Ala Lys Tyr Pro Asp His Ile Pro Asp Tyr Phe Lys
254 65          70          75          80
257 Gln Ser Phe Pro Lys Gly Phe Ser Trp Glu Arg Ser Leu Met Phe Glu
258          85          90          95

```

## RAW SEQUENCE LISTING

DATE: 03/03/2006

PATENT APPLICATION: US/10/525,365

TIME: 13:01:12

Input Set : A:\P26794.ST25.txt

Output Set: N:\CRF4\03032006\J525365.raw

261 Asp Gly Gly Val Cys Ile Ala Thr Asn Asp Ile Thr Leu Lys Gly Asp  
 262 100 105 110  
 265 Thr Phe Phe Asn Lys Val Arg Phe Asp Gly Val Asn Phe Pro Pro Asn  
 266 115 120 125  
 269 Gly Pro Val Met Gln Lys Lys Thr Leu Lys Trp Glu Ala Ser Thr Glu  
 270 130 135 140  
 273 Lys Met Tyr Leu Arg Asp Gly Val Leu Thr Gly Asp Ile Thr Met Ala  
 274 145 150 155 160  
 277 Leu Leu Leu Lys Gly Asp Val His Tyr Arg Cys Asp Phe Arg Thr Thr  
 278 165 170 175  
 281 Tyr Lys Ser Arg Gln Glu Gly Val Lys Leu Pro Gly Tyr His Phe Val  
 282 180 185 190  
 285 Asp His Cys Ile Ser Ile Leu Arg His Asp Lys Asp Tyr Asn Glu Val  
 286 195 200 205  
 289 Lys Leu Tyr Glu His Ala Val Ala His Ser Gly Leu Pro Asp Asn Val  
 290 210 215 220  
 293 Lys  
 294 225

297 &lt;210&gt; SEQ ID NO: 6

298 &lt;211&gt; LENGTH: 678

299 &lt;212&gt; TYPE: DNA

300 &lt;213&gt; ORGANISM: Trachyphyllia geoffroyi

302 &lt;400&gt; SEQUENCE: 6

303 atgagtctga ttaaaccaga aatgaagatc aagctgctta tggaaggcaa tgtaaaccggg 60  
 305 caccagtttg ttattgaggg agatggaaaa ggccatcctt ttgagggaaa acagagtatg 120  
 307 gacctttag tcaaagaagg cgcacctctc ccttttgcct acgatatctt gacaacagca 180  
 309 ttccattatg gtaacagggt ttttgctaaa taccagacc atataccaga ctacttcaag 240  
 311 cagtcgtttc ccaaagggtt ttcttgggag cgaagcctga tgttcgagga cgggggcggt 300  
 313 tgcacgcta caaatgacat aacactgaaa ggagacactt tttttaacaa agttcgattt 360  
 315 gatggcgtaa actttccccc aaatggctct gttatgcaga agaagactct gaaatgggag 420  
 317 gcatccactg agaaaatgta tttgcgtgat ggagtgttga cgggcgatat taccatggct 480  
 319 ctgctgctta aaggagatgt ccattaccga tgtgacttca gaactactta caaatctagg 540  
 321 caggaggggtg tcaagttgcc aggatatcac tttgtcgatc actgcatcag catattgagg 600  
 323 catgacaaag actacaacga ggtaagctg tatgagcatg ctgttgccca ttctggattg 660  
 325 cgggacaacg tcaagtaa 678

328 &lt;210&gt; SEQ ID NO: 7

329 &lt;211&gt; LENGTH: 225

330 &lt;212&gt; TYPE: PRT

331 &lt;213&gt; ORGANISM: Trachyphyllia geoffroyi

333 &lt;400&gt; SEQUENCE: 7

335 Met Ser Leu Ile Lys Pro Glu Met Lys Ile Lys Leu Leu Met Glu Gly  
 336 1 5 10 15  
 339 Asn Val Asn Gly His Gln Phe Val Ile Glu Gly Asp Gly Lys Gly His  
 340 20 25 30  
 343 Pro Phe Glu Gly Lys Gln Ser Met Asp Leu Val Val Lys Glu Gly Ala  
 344 35 40 45  
 347 Pro Leu Pro Phe Ala Tyr Asp Ile Leu Thr Thr Ala Phe His Tyr Gly  
 348 50 55 60  
 351 Asn Arg Val Phe Ala Lys Tyr Pro Asp His Ile Pro Asp Tyr Phe Lys

## RAW SEQUENCE LISTING

DATE: 03/03/2006

PATENT APPLICATION: US/10/525,365

TIME: 13:01:12

Input Set : A:\P26794.ST25.txt

Output Set: N:\CRF4\03032006\J525365.raw

```

352 65          70          75          80
355 Gln Ser Phe Pro Lys Gly Phe Ser Trp Glu Arg Ser Leu Met Phe Glu
356          85          90          95
359 Asp Gly Gly Val Cys Ile Ala Thr Asn Asp Ile Thr Leu Lys Gly Asp
360          100          105          110
363 Thr Phe Phe Asn Lys Val Arg Phe Asp Gly Val Asn Phe Pro Pro Asn
364          115          120          125
367 Gly Pro Val Met Gln Lys Lys Thr Leu Lys Trp Glu Ala Ser Thr Glu
368          130          135          140
371 Lys Met Tyr Leu Arg Asp Gly Val Leu Thr Gly Asp Ile Arg Met Glu
372 145          150          155          160
375 Leu Leu Leu Lys Gly Asp Val His Tyr Arg Cys Asp Phe Arg Thr Thr
376          165          170          175
379 Tyr Lys Ser Arg Gln Glu Gly Val Lys Leu Pro Gly Tyr His Phe Val
380          180          185          190
383 Asp His Cys Ile Ser Ile Leu Arg His Asp Lys Asp Tyr Asn Glu Val
384          195          200          205
387 Lys Leu Tyr Glu His Ala Val Ala His Ser Gly Leu Pro Asp Asn Val
388 210          215          220
391 Lys
392 225
395 <210> SEQ ID NO: 8
396 <211> LENGTH: 678
397 <212> TYPE: DNA
398 <213> ORGANISM: Trachyphyllia geoffroyi
400 <400> SEQUENCE: 8
401 atgagtctga ttaaaccaga aatgaagatc aagctgctta tggaaggcaa tgtaaacggg 60
403 caccagtttg ttattgaggg agatggaaaa ggccatcctt ttgagggaaa acagagtatg 120
405 gaccttgtag tcaaagaagg cgcacctctc ccttttgctt acgatatctt gacaacagca 180
407 ttccattatg gtaacagggt ttttgctaaa taccagacc atataccaga ctacttcaag 240
409 cagtcgtttc ccaaagggtt ttcttgggag cgaagcctga tgttcgagga cgggggcgtt 300
411 tgcacgcta caaatgacat aacactgaaa ggagacactt tttttaacaa agttcgattt 360
413 gatggcgtaa actttccccc aaatggtcct gttatgcaga agaagactct gaaatgggag 420
415 gcatccactg agaaaatgta ttgctgat ggagtgttga cgggcgatat taggatggag 480
417 ctgctgctta aaggagatgt ccattaccga tgtgacttca gaactactta caaatctagg 540
419 caggaggggtg tcaagttgcc aggatatcac tttgtcgatc actgcacag catattgagg 600
421 catgacaaag actacaacga ggttaagctg tatgagcatg ctgttgccca ttctggattg 660
423 ccggacaacg tcaagtaa 678
426 <210> SEQ ID NO: 9
427 <211> LENGTH: 229
428 <212> TYPE: PRT
429 <213> ORGANISM: Scolymia vitiensis
431 <400> SEQUENCE: 9
433 Met Val Ser Val Ile Lys Asp Glu Met Lys Val Asn Leu Arg Met Glu
434 1          5          10          15
437 Gly Ser Val Asn Gly His Asp Phe Val Ile Asp Gly Leu Gly Ser Gly
438          20          25          30
441 Lys Pro Lys Glu Gly Thr Gln Thr Ile Glu Leu Lys Val Val Lys Gly
442          35          40          45

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/525,365

DATE: 03/03/2006  
TIME: 13:01:13

Input Set : A:\P26794.ST25.txt  
Output Set: N:\CRF4\03032006\J525365.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:21; N Pos. 24,25,29,30,34,35

Seq#:27; N Pos. 7,9

**VERIFICATION SUMMARY**

DATE: 03/03/2006

PATENT APPLICATION: US/10/525,365

TIME: 13:01:13

Input Set : A:\P26794.ST25.txt

Output Set: N:\CRF4\03032006\J525365.raw

L:1004 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0

L:1102 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:0